

OIPE

3

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/874,390

DATE: 10/09/2001

TIME: 09:23:02

Input Set : A:\98017041.app

Output Set: N:\CRF3\10092001\I874390.raw

3 <110> APPLICANT: Clausen, Henrik
 5 <120> TITLE OF INVENTION: UDP-N-Acetylglucosamine:
 6 Galactose-beta-1,3-N-Acetylglactosamine-alpha-R /
 7 N-Acetylglucosamine-beta-1,3-N-Acetylglactosamine-alph
 8 a-R (GlcNAc to GalNAc)
 9 beta-1,6-N-Acetylglucosaminyltransferase, C2/4
 11 <130> FILE REFERENCE: P199801704 WO JNY
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/874,390
 C--> 14 <141> CURRENT FILING DATE: 2001-06-04
 16 <150> PRIOR APPLICATION NUMBER: DK PA 1998 01605
 17 <151> PRIOR FILING DATE: 1998-12-04
 19 <160> NUMBER OF SEQ ID NOS: 10
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 2319
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Homo sapiens
 28 <220> FEATURE:
 29 <221> NAME/KEY: CDS
 30 <222> LOCATION: (496)..(1809)
 31 <223> OTHER INFORMATION: cDNA sequence
 33 <400> SEQUENCE: 1
 34 attaaactggg ttttctatt tatctatcct ctgcattac ttctctgagt cagagcctct 60
 36 tctctctaag tcacgggaac tgcccttgct acttgtagacc tgccctttac tcagcagttt 120
 38 ttgttctggg aagccctggg attctgctaa tacctatcac tgtaggtgct gaagggaaac 180
 40 agatgaagaa catgacctca aggagcttcc tgtcaatgag aagaccaagc tgacgcctgg 240
 42 caaagatatt aaagaggagc ctgaaactgt tccttggaac tcttatgaat gtcagaaaat 300
 44 accttttgga gggttagaag atcaggggac atgggtgttc acatttgctg ccacggaaca 360
 46 ccgccagtct tcaactggaa acagaatcac gccttgtaga gagatcatcc ctaagcagga 420
 48 gagaagctac taaaggattg tgcctcctc caccttccct gtgctcggtc tccacctgtc 480
 50 tccattctg tgacg atg gtt caa tgg aag aga ctc tgc cag ctg cat tac 531
 51 Met Val Gln Trp Lys Arg Leu Cys Gln Leu His Tyr
 52 1 5 10
 54 ttg tgg gct ctg ggc tgc tat atg ctg ctg gcc act gtg gct ctg aaa 579
 55 Leu Trp Ala Leu Gly Cys Tyr Met Leu Leu Ala Thr Val Ala Leu Lys
 56 15 20 25
 58 ctt tct ttc agg ttg aag tgt gac tct gac cac ttg ggt ctg gag tcc 627
 59 Leu Ser Phe Arg Leu Lys Cys Asp Ser Asp His Leu Gly Leu Glu Ser
 60 30 35 40
 62 agg gaa tct caa agc cag tac tgt agg aat atc ttg tat aat ttc ctg 675
 63 Arg Glu Ser Gln Ser Gln Tyr Cys Arg Asn Ile Leu Tyr Asn Phe Leu
 64 45 50 55 60
 66 aaa ctt cca gca aag agg tct atc aac tgt tca ggg gtc acc cga ggg 723
 67 Lys Leu Pro Ala Lys Arg Ser Ile Asn Cys Ser Gly Val Thr Arg Gly
 68 65 70 75
 70 gac caa gag gca gtg ctt cag gct att ctg aat aac ctg gag gtc aag 771
 71 Asp Gln Glu Ala Val Leu Gln Ala Ile Leu Asn Asn Leu Glu Val Lys

ENTERED

RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/874,390

TIME: 09:23:02

Input Set : A:\98017041.app

Output Set: N:\CRF3\10092001\I874390.raw

72	80	85	90	
74 aag aag cga gag cct ttc aca gac acc cac tac ctc tcc ctc acc aga	819			
75 Lys Lys Arg Glu Pro Phe Thr Asp Thr His Tyr Leu Ser Leu Thr Arg				
76 95 100 105				
78 gac tgt gag cac ttc aag gct gaa agg aag ttc ata cag ttc cca ctg	867			
79 Asp Cys Glu His Phe Lys Ala Glu Arg Lys Phe Ile Gln Phe Pro Leu				
80 110 115 120				
82 agc aaa gaa gag gtg gag ttc cct att gca tac tct atg gtg att cat	915			
83 Ser Lys Glu Glu Val Glu Phe Pro Ile Ala Tyr Ser Met Val Ile His				
84 125 130 135 140				
86 gag aag att gaa aac ttt gaa agg cta ctg cga gct gtg tat gcc cct	963			
87 Glu Lys Ile Glu Asn Phe Glu Arg Leu Leu Arg Ala Val Tyr Ala Pro				
88 145 150 155				
90 cag aac ata tac tgt gtc cat gtg gat gag aag tcc cca gaa act ttc	1011			
91 Gln Asn Ile Tyr Cys Val His Val Asp Glu Lys Ser Pro Glu Thr Phe				
92 160 165 170				
94 aaa gag gcg gtc aaa gca att att tct tgc ttc cca aat gtc ttc ata	1059			
95 Lys Glu Ala Val Lys Ala Ile Ile Ser Cys Phe Pro Asn Val Phe Ile				
96 175 180 185				
98 gcc agt aag ctg gtt cgg gtg gtt tat gcc tcc tgg tcc agg gtg caa	1107			
99 Ala Ser Lys Leu Val Arg Val Val Tyr Ala Ser Trp Ser Arg Val Gln				
100 190 195 200				
102 gct gac ctc aac tgc atg gaa gac ttg ctc cag agc tca gtg ccg tgg	1155			
103 Ala Asp Leu Asn Cys Met Glu Asp Leu Leu Gln Ser Ser Val Pro Trp				
104 205 210 215 220				
106 aaa tac ttc ctg aat aca tgt ggg acg gac ttt cct ata aag agc aat	1203			
107 Lys Tyr Phe Leu Asn Thr Cys Gly Thr Asp Phe Pro Ile Lys Ser Asn				
108 225 230 235				
110 gca gag atg gtc cag gct ctc aag atg ttg aat ggg agg aat agc atg	1251			
111 Ala Glu Met Val Gln Ala Leu Lys Met Leu Asn Gly Arg Asn Ser Met				
112 240 245 250				
114 gag tca gag gta cct cct aag cac aaa gaa acc cgc tgg aaa tat cac	1299			
115 Glu Ser Glu Val Pro Pro Lys His Lys Glu Thr Arg Trp Lys Tyr His				
116 255 260 265				
118 ttt gag gta gtg aga gac aca tta cac cta acc aac aag aag aag gat	1347			
119 Phe Glu Val Val Arg Asp Thr Leu His Leu Thr Asn Lys Lys Lys Asp				
120 270 275 280				
122 cct ccc cct tat aat tta act atg ttt aca ggg aat gcg tac att gtg	1395			
123 Pro Pro Pro Tyr Asn Leu Thr Met Phe Thr Gly Asn Ala Tyr Ile Val				
124 285 290 295 300				
126 gct tcc cga gat ttc gtc caa cat gtt ttg aag aac cct aaa tcc caa	1443			
127 Ala Ser Arg Asp Phe Val Gln His Val Leu Lys Asn Pro Lys Ser Gln				
128 305 310 315				
130 caa ctg att gaa tgg gta aaa gac act tat agc cca gat gaa cac ctc	1491			
131 Gln Leu Ile Glu Trp Val Lys Asp Thr Tyr Ser Pro Asp Glu His Leu				
132 320 325 330				
134 tgg gcc acc ctt cag cgt gca cgg tgg atg cct gcc tct gtt ccc aac	1539			
135 Trp Ala Thr Leu Gln Arg Ala Arg Trp Met Pro Gly Ser Val Pro Asn				
136 335 340 345				

RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/874,390

TIME: 09:23:02

Input Set : A:\98017041.app

Output Set: N:\CRF3\10092001\I874390.raw

```

138 cac ccc aag tac gac atc tca gac atg act tct att gcc agg ctg gtc 1587
139 His Pro Lys Tyr Asp Ile Ser Asp Met Thr Ser Ile Ala Arg Leu Val
140      350      355      360
142 aag tgg cag ggt cat gag gga gac atc gat aag ggt gct cct tat gct 1635
143 Lys Trp Gln Gly His Glu Gly Asp Ile Asp Lys Gly Ala Pro Tyr Ala
144 365      370      375      380
146 ccc tgc tct gga atc cac cag cgg gct atc tgc gtt tat ggg gct ggg 1683
147 Pro Cys Ser Gly Ile His Gln Arg Ala Ile Cys Val Tyr Gly Ala Gly
148      385      390      395
150 gac ttg aat tgg atg ctt caa aac cat cac ctg ttg gcc aac aag ttt 1731
151 Asp Leu Asn Trp Met Leu Gln Asn His His Leu Leu Ala Asn Lys Phe
152      400      405      410
154 gac cca aag gta gat gat aat gct ctt cag tgc tta gaa gaa tac cta 1779
155 Asp Pro Lys Val Asp Asp Asn Ala Leu Gln Cys Leu Glu Glu Tyr Leu
156      415      420      425
158 cgt tat aag gcc atc tat ggg act gaa ctt tgagacacac tatgagagcg 1829
159 Arg Tyr Lys Ala Ile Tyr Gly Thr Glu Leu
160      430      435
162 ttgctacctg tggggcaaga gcatgtacaa acatgctcag aacttgctgg gacagtgtgg 1889
164 gtgggagacc agggctttgc aattcgtggc atcctttagg ataagagggc tgctattaga 1949
166 ttgtgggtaa gtagatcttt tgccttgcaa attgctgcct gggatgaatgc tgcttggtct 2009
168 ctaccccta accctagtag ttcctccact aactttctca ctaagtgaga atgagaactg 2069
170 ctgtgatagg gagagtgaag gagggatatg tggtagagca cttgatttca gttgaatgcc 2129
172 tgctggtagc ttttccattc tgtggagctg ccgttcctaa taattccagg tttggtagcg 2189
174 tggaggagaa ctttgatgga aagagaacct tcccttctgt actgttaact taaaaataaa 2249
176 tagctcctga ttcaaagtat tacctctact ttttgcttag tatgccagaa ataataataa 2309
178 tctaaacaga 2319
181 <210> SEQ ID NO: 2
182 <211> LENGTH: 438
183 <212> TYPE: PRT
184 <213> ORGANISM: Homo sapiens
186 <400> SEQUENCE: 2
187 Met Val Gln Trp Lys Arg Leu Cys Gln Leu His Tyr Leu Trp Ala Leu
188 1      5      10      15
190 Gly Cys Tyr Met Leu Leu Ala Thr Val Ala Leu Lys Leu Ser Phe Arg
191      20      25      30
193 Leu Lys Cys Asp Ser Asp His Leu Gly Leu Glu Ser Arg Glu Ser Gln
194      35      40      45
196 Ser Gln Tyr Cys Arg Asn Ile Leu Tyr Asn Phe Leu Lys Leu Pro Ala
197      50      55      60
199 Lys Arg Ser Ile Asn Cys Ser Gly Val Thr Arg Gly Asp Gln Glu Ala
200 65      70      75      80
202 Val Leu Gln Ala Ile Leu Asn Asn Leu Glu Val Lys Lys Lys Arg Glu
203      85      90      95
205 Pro Phe Thr Asp Thr His Tyr Leu Ser Leu Thr Arg Asp Cys Glu His
206      100      105      110
208 Phe Lys Ala Glu Arg Lys Phe Ile Gln Phe Pro Leu Ser Lys Glu Glu
209      115      120      125
211 Val Glu Phe Pro Ile Ala Tyr Ser Met Val Ile His Glu Lys Ile Glu

```

RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/874,390

TIME: 09:23:02

Input Set : A:\98017041.app

Output Set: N:\CRF3\10092001\I874390.raw

```

212      130      135      140
214 Asn Phe Glu Arg Leu Leu Arg Ala Val Tyr Ala Pro Gln Asn Ile Tyr
215 145      150      155      160
217 Cys Val His Val Asp Glu Lys Ser Pro Glu Thr Phe Lys Glu Ala Val
218      165      170      175
220 Lys Ala Ile Ile Ser Cys Phe Pro Asn Val Phe Ile Ala Ser Lys Leu
221      180      185      190
223 Val Arg Val Val Tyr Ala Ser Trp Ser Arg Val Gln Ala Asp Leu Asn
224      195      200      205
226 Cys Met Glu Asp Leu Leu Gln Ser Ser Val Pro Trp Lys Tyr Phe Leu
227      210      215      220
229 Asn Thr Cys Gly Thr Asp Phe Pro Ile Lys Ser Asn Ala Glu Met Val
230 225      230      235      240
232 Gln Ala Leu Lys Met Leu Asn Gly Arg Asn Ser Met Glu Ser Glu Val
233      245      250      255
235 Pro Pro Lys His Lys Glu Thr Arg Trp Lys Tyr His Phe Glu Val Val
236      260      265      270
238 Arg Asp Thr Leu His Leu Thr Asn Lys Lys Lys Asp Pro Pro Pro Tyr
239      275      280      285
241 Asn Leu Thr Met Phe Thr Gly Asn Ala Tyr Ile Val Ala Ser Arg Asp
242      290      295      300
244 Phe Val Gln His Val Leu Lys Asn Pro Lys Ser Gln Gln Leu Ile Glu
245 305      310      315      320
247 Trp Val Lys Asp Thr Tyr Ser Pro Asp Glu His Leu Trp Ala Thr Leu
248      325      330      335
250 Gln Arg Ala Arg Trp Met Pro Gly Ser Val Pro Asn His Pro Lys Tyr
251      340      345      350
253 Asp Ile Ser Asp Met Thr Ser Ile Ala Arg Leu Val Lys Trp Gln Gly
254      355      360      365
256 His Glu Gly Asp Ile Asp Lys Gly Ala Pro Tyr Ala Pro Cys Ser Gly
257      370      375      380
259 Ile His Gln Arg Ala Ile Cys Val Tyr Gly Ala Gly Asp Leu Asn Trp
260 385      390      395      400
262 Met Leu Gln Asn His His Leu Leu Ala Asn Lys Phe Asp Pro Lys Val
263      405      410      415
265 Asp Asp Asn Ala Leu Gln Cys Leu Glu Glu Tyr Leu Arg Tyr Lys Ala
266      420      425      430
268 Ile Tyr Gly Thr Glu Leu
269      435

```

273 <210> SEQ ID NO: 3

274 <211> LENGTH: 21

275 <212> TYPE: DNA

276 <213> ORGANISM: Artificial Sequence

278 <220> FEATURE:

279 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer

281 <400> SEQUENCE: 3

282 ggaagttcat acagttccca c

21

285 <210> SEQ ID NO: 4

286 <211> LENGTH: 21

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/874,390

DATE: 10/09/2001

TIME: 09:23:02

Input Set : A:\98017041.app

Output Set: N:\CRF3\10092001\I874390.raw

```

287 <212> TYPE: DNA
288 <213> ORGANISM: Artificial Sequence
290 <220> FEATURE:
291 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
293 <400> SEQUENCE: 4
294 cctcccatc aacatcttga g 21
297 <210> SEQ ID NO: 5
298 <211> LENGTH: 31
299 <212> TYPE: DNA
300 <213> ORGANISM: Artificial Sequence
302 <220> FEATURE:
303 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
305 <400> SEQUENCE: 5
306 agcgaattca gctcaaagtt cagtcccata g 31
309 <210> SEQ ID NO: 6
310 <211> LENGTH: 21
311 <212> TYPE: DNA
312 <213> ORGANISM: Artificial Sequence
314 <220> FEATURE:
315 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
317 <400> SEQUENCE: 6
318 gtgggaactg tatgaacttc c 21
321 <210> SEQ ID NO: 7
322 <211> LENGTH: 30
323 <212> TYPE: DNA
324 <213> ORGANISM: Artificial Sequence
326 <220> FEATURE:
327 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
329 <400> SEQUENCE: 7
330 gcagaattca tggttcaatg gaagagactc 30
333 <210> SEQ ID NO: 8
334 <211> LENGTH: 26
335 <212> TYPE: DNA
336 <213> ORGANISM: Artificial Sequence
338 <220> FEATURE:
339 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
341 <400> SEQUENCE: 8
342 cgagaattca ggttgaagtg tgactc 26
345 <210> SEQ ID NO: 9
346 <211> LENGTH: 21
347 <212> TYPE: DNA
348 <213> ORGANISM: Artificial Sequence
350 <220> FEATURE:
351 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
353 <400> SEQUENCE: 9
354 gctcgtctc cacctgtctc c 21
357 <210> SEQ ID NO: 10
358 <211> LENGTH: 21
359 <212> TYPE: DNA

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/874,390

DATE: 10/09/2001

TIME: 09:23:03

Input Set : A:\98017041.app

Output Set: N:\CRF3\10092001\I874390.raw

L:13 M:270 C: Current Application Number differs, Replaced Application Number

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date